Motion Imagery Standards Board **Frequently Asked Questions**	**Updated 09 Feb 2011**

General

What is the Motion Imagery Standards Board (MISB)?
Why should I care about the MISB? Where and when do MISB requirements apply?
What constitutes a motion imagery system (as defined by the MISB)?
What is the difference between motion imagery and full motion video?
I'm building a new motion imagery system. What is the quick-and-dirty on what I have to do to be MISB compliant?
I'm building a new motion imagery system. What I should avoid doing?
Where is the MISB website, and what can I find there?
What is the difference between an EG, an RP and a Standard?
I need a viewer to play the video and metadata. What should I choose?
I want to get my MI system certified. How do I start the process?

Compression

Of the three approved compression algorithms (MPEG-2, H.264 and JPEG 2000), which one should I use?
If I have a choice between MPEG-2 and H.264, which one is recommended?
Why are codec's like Microsoft's VC-1 not adopted by the MISB?

Metadata

What's wrong with *MISB EG 0104*? You guys promulgated it, after all.
What is KLV metadata?
KLV metadata isn't human readable! Why?
Where do I find definitions for KLV keys?
Why are there two metadata dictionaries?
Which metadata dictionary (MISB or SMPTE) has precedence?
How can I tell if a key is in SMPTE RP 210 or MISB Standard 0807?
If I need new keys registered, should I go to SMPTE or the MISB?
What KLV metadata do I need to use?

Time Stamping and Synchronization

What is the difference between asynchronous and synchronous metadata?
What is time stamping?
I don't understand time stamping, why do I need it?

Transport Protocols and File Wrappers

Can MPEG-2 Transport Stream carry H.264 encoded content?
Why is MPEG-2 Transport a desired carrier for motion imagery?
Why can't I use other transport protocols like ASF that work naturally with Windows products?
What is RTP/RTCP/RTSP used for?

Large Volume Streaming Data (LVSD)

<u>General</u>

Q1. What is the Motion Imagery Standards Board (MISB)?
A1. The Motion Imagery Standards Board (MISB) was established in accordance with DoD Directive 5105.60 "to formulate, review, and recommend standards for motion imagery, associated metadata, audio, and other related systems" for the Department of Defense (DoD), Intelligence Community (IC), and National System for Geospatial-Intelligence (NSG). The MISB exists under the Geospatial Intelligence Standards Working Group (GWG) of the National Center for Geospatial Intelligence Standards.

The MISB meets three times a year (typically May, September and December) in the Washington D.C. metropolitan area. Each meeting runs approximately one week in length. The MISB is comprised of seven working groups that address different functional areas regarding motion imagery. For a description of the various working groups and what functional areas they address consult the "Surfing the MISP" document which may be found on the MISB website (www.gwg.nga.mil/misb).

Q2. Why should I care about the MISB? Where and when do MISB requirements apply?
A2. Any motion imagery/full motion video (MI/FMV) system subject to the DoD IT Standards Registry (DISR) (formerly the DoD Joint Technical Architecture) and/or the NSG Technical Architecture is subject to MISB standards and requirements. If you are manufacturing motion imagery systems or components for use within the DoD/IC communities, those systems and components are subject to MISB standards and requirements.

Q3. What constitutes a motion imagery system (as defined by the MISB)?
A3. Any imaging system that collects at a rate of one frame per second (1 Hz) or faster, over a common field of regard, is a motion imagery system. This explicitly includes, but is not limited to, Electro-optical (EO), Infrared (IR), Multi-spectral (MSI), and Hyper-spectral (HSI) systems. Video Teleconference (VTC), Video Telemedicine, and Video Support Services applications **DO NOT** fall within the purview of the MISB and are not subject to its requirements.

Q4. What is the difference between motion imagery and full motion video?
A4. The MISB makes no formal distinction between the terms "motion imagery" and "full motion video (FMV)." However, motion imagery must contain metadata. Some entities call video with no metadata "full motion video". Historically, however, FMV has been that subset of motion imagery at television-like frame rates (24 – 60 Hz).

Q5. I'm building a new motion imagery system. What is the quick-and-dirty on what I have to do to be MISB compliant?
A5. To be MISB compliant, any new motion imagery system must:
1) Be digital
2) Produce a compliant MPEG-2 Transport Stream (TS). Note this does not apply to JPEG 2000 based systems or streaming (RTP) based systems.
3) Use MPEG-2, MPEG-4 Part 10 (H.264/AVC), or JPEG 2000 image compression
4) Produce non-destructive (not "burned in") metadata
5) Comply with MISB *Standards 0601*, *0102*, and *0604*

Older MISB compliant systems used *EG 0104*, which has been deprecated. The Motion Imagery Standards Profile (MISP) codifies all MISB requirements, Standards, Recommended Practices, and Engineering Guidelines. The MISP is found on the MISB website, and is in the DISR.

Q6. I'm building a new motion imagery system. What I should avoid doing?

A6. **Do not** build:
1) Analog systems
2) Digital systems that use interlaced scanning
3) Destructive ("burned in") metadata
4) MISB EG 0104
5) Systems that utilize proprietary file formats, metadata encodings or compression algorithms
6) Systems that utilize standardized file formats, metadata encodings and compression algorithms not covered by the MISP. Just because a standard exists, that does not mean it has been adopted into the MISP.

Q7. Where is the MISB website, and what can I find there?

A7. The MISB website is www.gwg.nga.mil/misb . The MISP (Motion Imagery Standards Profile) and all current Standards, Recommended Practices (RPs), and Engineering Guidelines (EGs) can be found there. A good starting point is to review the MISP, which includes references to all subsequent MISB Standards, RPs, and EGs. If you need access to draft documents, test files, and other support documentation follow the instructions on the website to apply for an account to access the MISB protected website.

Q8. What is the difference between an EG, an RP and a Standard?

A8. The MISB recommends that implementers adhere to all EGs, RPs and Standards that it publishes. However, we realize that special circumstances and needs arise that may prevent this. Engineering Guidelines provide guidance that implementers should follow whenever possible. They may be viewed as a particular solution to a problem and frequently represent the initial attempt at solving a problem.

As an EG evolves and becomes standard practice in the community, it may be turned into an RP or even a Standard. It may also be superseded by subsequent EGs, RPs and Standards. As an example we may consider EG 0104, the Predator UAV Basic Universal Metadata Set. This EG started out as a way to take the Closed Caption metadata out of predator analog video feeds and encode it using KLV. It is clearly acknowledged in the EG that this was a temporary solution to ease the transition from analog to digital video and that subsequent standards would define a more complete metadata set. This metadata set, defined in Standard 0601, the *UAS Datalink Local Metadata Set*, has subsequently replaced EG 0104 for all new UAS systems.

Recommend Practices are typically profiles of a Standard. They represent implementation practices that the MISB strongly recommends all implementers and programs should follow. They carry more weight than EGs and RPs typically have been validated in the field as a best practice. Programs and implementers may elect not to follow RPs, but this must not be done without good justification and consent of the MISB. As an example, consider RP 0705 the LVSD Compression Profile. This document defines a compression profile for JPEG 2000 to be used by LVSD systems. It profiles the ISO JPEG 2000 document (ISO/IEC 15444-1:2004). RPs may be superseded by

subsequent RPs and Standards as standard practices evolve.

A Standard represents a requirement that all MISP compliant systems shall comply with. Standard 0601 is the local data set that all UAS systems must comply with. Over time RPs and even EGs may develop into Standards. Implementers and programs are required to comply with all relevant Standards.

Q9. **I need a viewer to play the video and metadata. What should I choose?**

A9. There are several tools available from companies like General Dynamics, SAIC, and PAR Government Systems. The MISB cannot make recommendations regarding software and hardware solutions.

Q10. **I want to get my MI system certified. How do I start the process?**

A10. The Joint Interoperability Test Command (JITC) has the authority to test and certify MI systems for compliance to MISB standards. Contact the Motion Imagery Standards Laboratory (MIS-LAB) at 2001 Brainard Road, Bldg 57305 Fort Huachuca, AZ 85613. The JITC's Motion Imagery Standards Laboratory (MIS-LAB) website may be found at (http://jitc.fhu.disa.mil/mis/index.html). To schedule compliance testing contact the MIS-LAB.

Compression

Q11. **Of the three approved compression algorithms (MPEG-2, H.264 and JPEG 2000), which one should I use?**

A11. The correct answer depends on your systems. H.264 yields the best quality for low bandwidth applications. For similar video quality MPEG-2 compression needs roughly twice the bandwidth. H.264 is quickly replacing MPEG-2 in the commercial world.

JPEG 2000 is intraframe compression rather than intra- and interframe as is found in MPEG-2 and H.264. JPEG 2000 therefore consumes 2 – 3 times the bandwidth of MPEG-2. However, JPEG-2000 accommodates very large frame (Gpixels) sizes and has low (1 frame) latency. JPEG 2000 has features that make it very useful in large volume steaming data (LVSD) applications, while H.264 and MPEG-2 are typically used in full motion video applications. For more information regarding LVSD systems see the section below.

Q12. **If I have a choice between MPEG-2 and H.264, which one is recommended?**

A12. H.264 offers better performance over MPEG-2 in reduced bandwidth for similar quality (about 2 to 1). This improvement comes with increased complexity in the encoder and decoder, which affects overall cost, but with the rapid adoption of H.264 in the commercial world this is becoming less of an issue. DISA is pushing forward H.264 over MPEG-2 throughout their networks. For HD applications, H.264 offers the best performance overall.

Q13. **Why are codec's like Microsoft's VC-1 not adopted by the MISB?**

A13. The goal of the MISB is to promote technologies that are standards-based in order to ensure interoperable solutions across the PED process. Although SMPTE has recently adopted Microsoft's VC-1 video codec as a standard, VC-1 does not offer a significant performance advantage over H.264. Adding more codec standards to the MISB suite of technologies can create interoperability issues, so technologies are carefully reviewed for their added value.

Q14. What's wrong with *MISB EG 0104*? The MISB promulgated it, after all.

A14. *MISB EG 0104: Predator UAV Basic Universal Metadata Set* was the first step in moving away from the analog metadata used by the initial RQ-1's. We've learned a great deal over the past decade-plus and we now know that EG 0104 falls short in a number of regards. Although still supported by the MISB for legacy systems, there is no reason to use it in a new system. Any information you can convey with *EG 0104* you can convey with *Standard 0601* with greater precision and bit-efficiency.

Q15. What is KLV metadata?

A15. KLV stands for Key-Length-Value. KLV metadata comes in self-contained binary units. The Key tells you what the metadata element describes, the Length how long in bytes the value of the data, and the Value contains the actual data. KLV metadata is very bit–efficient. The Society of Motion Picture and Television Engineers (SMPTE) standard, *SMPTE 336M: Data Encoding Protocol Using Key-Length-Value*, defines the KLV data encoding protocol.

Q16. KLV metadata isn't human readable! Why?

A16. That is a feature, not a bug. KLV is expressed in binary bits, which provide a very efficient representation of data. There is a great deal of padding in XML to make it "human readable" that wastes precious bandwidth. KLV metadata *can* be translated into human-readable XML (and *vice versa*) without loss of information, if necessary.

Q17. Where do I find definitions for KLV keys?

A17. The structure of KLV metadata is defined in *SMPTE 336M*. The actual metadata dictionaries are *SMPTE RP 210* and *MISB Standard 0807*.

Q18. Why are there two metadata dictionaries?

A18. SMPTE created the standard for KLV encoding of metadata. SMPTE produces and maintains a KLV metadata dictionary (*SMPTE RP 210*). Various organizations are allowed to buy part of the KLV domain name-space to maintain private metadata dictionaries. The DoD was the first organization to take advantage of this offer. Initially, most of the metadata keys used by the MISB were registered in *SMPTE RP 210*, but, over time, several issues became apparent. First, it can take 12 – 24 months to get a new KLV metadata key approved by SMPTE. Second, SMPTE does not give tight definitions to their metadata elements. MISB *Standard 0807* is the metadata dictionary for elements in the DoD private domain space. The MISB can assign keys quickly if necessary (a week is common), and can define their meaning and usage to whatever exactitude is necessary. Finally, because the keys in *Standard 0807* are not published to the general public, it is possible to maintain classified keys.

Q19. Which metadata dictionary (MISB or SMPTE) has precedence?

A19. MISB Standard 0807 has precedence over SMPTE RP 210.

Q20. How can I tell if a key is in SMPTE RP 210 or MISB Standard 0807?

A20. All KLV Keys are 16 bytes long. All SMPTE keys (including the DoD private keys in *Standard 0807*) begin with the 4-byte sequence 06 0E 2B 34 (in hexadecimal). Keys from *MISB Standard 0807* have the ninth byte set to 0E and the tenth byte set to 01, 02, or 03. A MISB key will, therefore, have the form 06 0E 2B 34 xx xx xx xx 0E [01, 02, or 03] xx xx xx xx xx xx. As a general rule, older MISB documents have *SMPTE RP 210* keys, and newer MISB documents have their keys registered in

MISB Standard 0807.

Q21. If I need new keys registered, should I go to SMPTE or the MISB?
A21. Go to the MISB. We can create keys faster and define their usage unambiguously.

Q22. What KLV metadata do I need to use?
A22. You need to support *Standard 0601: UAS Datalink Local Data Set* and *Standard 0102: Security Metadata Universal and Local Data Sets*. At a minimum, you must support the elements from those two sets called out by *MISB EG 0902: Motion Imagery Sensor Minimum Metadata Set*. Depending on your mission requirements and CONOPS, you may need to support more than the baseline elements from those two Standards or metadata defined in other MISB documents.

Time Stamping and Synchronization

Q23. What is the difference between asynchronous and synchronous metadata?
A23. In general, asynchronous metadata is collected at the point of acquisition without regard to accurate alignment to a particular frame in the motion imagery. Units of metadata travel in close proximity to corresponding events in the video, but this proximity can vary depending on how the MI and metadata information is processed. If the asynchronous metadata has time stamp information associated with it, the metadata can be correlated with the video frames (some interpolation of the metadata may also be required).

Synchronous metadata is collected in temporal alignment with the video stream. It is time stamped and associated to the imagery in a defined manner. Events in the imagery can then be accurately associated with the corresponding metadata. It is preferred that all future MI systems employ synchronous metadata.

Q24. What is time stamping?
A24. *All* metadata and video sources should be time stamped. Time stamps are as simple as a date and time record of when the metadata or a video frame is generated. The MISB has standardized on the types of date/time formats it supports and how this information should be placed within a video stream. *MISB RP 0103: Timing Reconciliation Universal Metadata Set for Digital Motion Imagery*, defines the KLV metadata element for time stamping metadata. *MISB RP 0603: Common Time Reference for Digital Motion Imagery Using Coordinated Universal Time (UTC)*, describes the usage of time stamps within the MISB guidelines and defines UTC as the preferred time stamp. *MISB RP 0605: Inserting Time Code and Metadata in High Definition Uncompressed Video,* describes how to insert time stamps into uncompressed motion imagery and *MISB Standard 0604: Time Stamping Compressed Motion Imagery* describes the insertion of time stamps into compressed motion imagery.

Q25. I don't understand time stamping, why do I need it?
A25. Time stamping aides in search and retrieval operations once motion imagery is archived. Time stamps help accurately align metadata with collected motion imagery for further event analysis and exploitation. It is not uncommon for platform metadata to be collected asynchronously relative to the motion imagery. For example, platform elevation, heading and speed might be collected at 7 Hz, while the motion imagery might be collected at 30 Hz. Clearly the metadata will not temporally align with the video frames except possibly once a second. Time stamping will allow for interpolation of the metadata, if needed, for processing or exploiting a given video frame.

Transport Protocols and File Wrappers

Q26. Can MPEG-2 Transport Stream carry H.264 encoded content?
A26. Yes. Updates to ISO standards have included this capability, and the MISB calls this Xon2 where "X" is either MPEG-2 or H.264 encoded motion imagery transported over MPEG-2 transport stream.

Q27. Why is MPEG-2 Transport Stream a desired carrier for motion imagery?
A27. MPEG-2 Transport Stream (TS) is a mature technology designed originally for digital television transmission. As a standard it is widely supported and many tools are available for testing and compliance. The value in motion imagery is greatly increased when augmented with metadata, and MPEG-2 TS provided an excellent technology to wrap the motion imagery with its metadata and maintain the combination as a unified package for subsequent use.

Q28. Why can't I use other transport protocols like ASF that work naturally with Windows products?
A28. ASF is a proprietary format controlled by a single company. Proprietary formats are not easily extended and there is typically little documentation regarding their structure and use. Third parties who wish to develop against proprietary formats are dependant upon the holder of the format to not introduce unexpected changes. Proprietary formats therefore do not promote interoperability. In addition, ASF has no facility to carry metadata.

Q29. What is RTP/RTCP/RTSP used for?
A29. Real-time Transport Protocol (RTP) is designed to deliver real time media, such as video and audio, over internet protocol links. Specifically, RTP addresses the public internet, where quality-of-service (QoS) is not guaranteed. RTP is a protocol layer added (typically) on top UDP that adds a time stamp and count to every data packet to aid the receiver in reconstructing the stream when packets suffer latency, become reordered, or are lost in the network. MPEG-2 transport stream does not do as well in such environments because it was designed for constant delay networks like broadcast. Some systems do use RTP to carry MPEG2 transport stream at the expense of additional data overhead and may be less robust in the presence of lost packets.

RTP generally is accompanied with the bi-directional server/client protocol RTCP (RTP Control Protocol). RTCP provides network and timing information between video senders (servers) and receivers (clients). Clients and servers use this information to determine Quality of Service (QoS) operating points and to maintain real-word time synchronization. Finally, RTSP (Real Time Streaming Protocol) provides information that allows clients and servers to describe and establish video streaming sessions and it gives clients TiVo-like control for the client to record, rewind, stop, play, and fast-forward the stream. The MISB is currently evaluating these technologies for future adoption.

Q30. What is JPIP used for?
A30. JPIP (JPEG 2000 Interactive Protocol) is similar in spirit to RTP and RTSP (there currently is no RTCP equivalent within JPIP). JPIP is a client/server streaming protocol that provides interactive delivery of JPEG 2000 compressed imagery. It allows a client to specify a region of interest out of a large image, at a desired resolution and image quality and have the data streamed to the client. Using JPEG 2000 and JPIP together it is possible to browse very large images (1 Gpixel and up) on lightweight clients (PDAs). This is possible because only small portions of the compressed image are streamed from the server to the client. As the client changes their viewing region, the server

streams new information to the client to update the image display. The MISB anticipates that JPIP will be very useful for LVSD applications (see below).

Q31. What are the differences between file transfer, progressive download, and streaming?

A31. File transfer is based on FTP, which is a protocol that guarantees complete delivery of the file to a receiver. FTP operates over TCP/IP, and therefore all packets are assured they will be received as transmitted. Because of this the download of a file using FTP can take a long time, and the user must wait for the content to be delivered in its entirety prior to viewing. Progressive download helps this by invoking a buffer in the receiver that will display the content after sufficient data has been received; the user must still wait, however.

Streaming is designed to accommodate real time delivery of content and is appropriate for live events and time-critical applications. Streaming operates over UDP/IP, and for this reason cannot guarantee that all packets transmitted will be received. The quality of content received via streaming may be low as the server/client attempt to deliver the stream as fast as possible to meet real time delivery. Image size, frame rate, and bits assigned to preserve image detail all may be adjusted to meet the channel bandwidth. See MISB *EG 0803 Delivery of Low Bandwidth Motion Imagery* and *TRM 07A Low Bandwidth Motion Imagery – Technologies* (protected site) for more information.

Q32. What is the best format to store/archive motion imagery?

A32. At this time, the MISB advocates the MPEG-2 Transport Stream (TS), AAF (Advanced Authoring Format) and MXF (Material eXchange Format) as file wrappers. MPEG-2 TS is a delivery format that also serves as container. AAF can accommodate historical editing and updates of content as it moves through its production. MXF is emerging as a format that can manage complex content and metadata, and is also designed for exchange of motion imagery. The "best" format to choose is application dependent.

Large Volume Streaming Data (LVSD)

Q33. What is Large Volume Streaming Data (LVSD)?

A33. LVSD is a NATO designation for a certain class of motion imagery sensors. In the U.S., the terms WALF (Wide Area Large Format), WAMI (Wide Area Motion Imagery), WAPS (Wide Area Persistent Surveillance) and WAAS (Wide Area Aerial Surveillance) are frequently used. All of these acronyms are used to describe essentially the same systems, so don't let the various terms confuse you. LVSD systems typically collect very large frame imagery (100 Mpixels to a 10 Gpixels per frame) using arrays of smaller cameras or multiple focal plane sensors which are then processed to form a large single image mosaic. The large image frames are collected at rates of 1 Hz or faster, therefore we treat them as motion imagery.

LVSD systems may incorporate more than one sensor modality. For example, an HD FMV (i.e. 24 – 60 Hz frame rate) camera might be present in addition to the large image array. LVSD systems typically collect very large volumes of data (terabytes to petabytes during a collection) and may provide motion imagery streaming services off of the platform during data collects. The MISB is currently working with systems under development and our NATO partners to define appropriate compression and metadata recommendations for these systems.

Q34. What standards are available for LVSD systems?

A34. LVSD systems will need to employ several different standards to fully realize their potential. LVSD systems that employ HD or SD FMV cameras should make use of H.264 and MPEG-2 TS for compression and formats for these sensors. The *MISB EG 0601* and *Standard 0102* data sets should be used for metadata. Implementers may also need to consider *MISB EG 0801: Profile 1: Photogrammetry Metadata Set for Digital Motion Imagery*. The MISB is currently working on a metadata set for LVSD systems; see *MISB EG 0810: Profile 2: KLV for LVSD Applications*. The proper choice of metadata set is really a function of system CONOPS. Implementers and system designers are strongly encouraged to think beyond the current limited use of their LVSD data. Other people will find new and interesting ways to exploit this motion imagery and appropriate metadata will be a necessity.

The large image frames generated by LVSD camera arrays are currently compressed on a frame by frame basis. All known systems are either using JPEG 2000 or plan to migrate to it. The MISB has developed a JPEG 2000 compression profile, *MISB RP 0705: LVSD Compression Profile* and a JPIP (JPEG 2000 Interactive Protocol) profile, *MISB RP 0811: JPIP Profile (Client/Server Functions)*. JPIP is a client/server interactive streaming protocol for delivery of JPEG 2000 compressed imagery.

The MISB is also developing file formats for LVSD systems. LVSD systems comprise large volumes of motion imagery data, metadata and may contain multiple sensor types. Most file formats are not flexible enough to serve as an LVSD container. The commercial video industry has very similar problems at the studio level. The need to contain metadata and video from disparate sources and maintain an accurate temporal relationship between all of the data lead to the development of the AAF and MXF (Material eXchange Format) standards. MXF is a SMPTE standard (see SMPTE 377M) and the AAF standard is maintained by the Advanced Media Workflow Association (see http://www.aafassociation.org/). The MISB has developed an AAF profile that is appropriate for LVSD the relevant document is *MISB RP 0301.3a: MISB Profile for Aerial Surveillance and Photogrammetry Applications (ASPA)*.

As new sensor modalities are added to LVSD systems and LVSD system CONOPS change, new standards, recommended practices and engineering guidelines will need to be developed. The MISB is seeking active participation from all programs and implementers in this area.

Q35. Wait a minute; you said that JPEG 2000 compression was not as good as MPEG-2 or H.264 for motion imagery. Why are you using it for LVSD?

A35. The reasons for this choice are many. JPEG 2000 provides a multi-resolution representation of the compressed image. This is very important when dealing with 100 Mpixel – 10 Gpixel images where it is impossible to view the full image at full resolution. When dealing with imagery of this size you typically look at the full image at reduced resolution and zoom in to increased resolutions as you define your areas of interest. JPEG 2000 excels at this. It is trivial to extract reduced resolution data sets (RRDS) from a JPEG 2000 compressed image. JPEG 2000 also provides easy region of interest access and decoding and the ability to adjust decoded/transmitted image quality on the fly. This allows users to select a desired spatial region of interest in a large image and even control the visual quality received.

Furthermore, JPEG 2000 allows for images with bit depths up to 32 bits/pixel/color component and it allows up to 16,000 color components, so it is suited for multi-spectral image compression. JPEG 2000 can even compress an image losslessly when needed and extract from the losslessly compressed image a reduced quality version to save on transmission bandwidth. The JPIP protocol

allows for very efficient transmission of portions of large compressed images over modest bandwidth links. Most of these features are simply not available within the MPEG-2 and H.264 standards.

The MISB recognizes that increasing the compression efficiency of JPEG 2000 by adding temporal compression capability to the standard would be advantageous for LVSD applications. We are currently examining the feasibility of doing this with our academic and government partners.